Slide 1 of 14

Title Slide: ELS:2002 Sample Design, Weights, Variance, and Missing Data

Slide 2 of 14

ELS:2002 used a complex design for sample selection. As discussed in the common modules, when analyzing data from complex sample surveys, certain procedures must be used to assure that estimates made from the data are representative of the population and that hypothesis tests are accurate. This module provides information about the weights and variables associated with estimates from ELS:2002 data.

Specifically, this module will present information about the ELS survey design to facilitate a more detailed discussion of sampling weights, standard error calculation, and missing and imputed data. This module will also describe the types of weights within ELS:2002 and provide guidance regarding how to select the appropriate weight for specific analyses.

For information about weighting and calculating appropriate standard errors more generally, please view the relevant common modules, if you have not already done so, which are accessible by clicking on the underlined screen text on this slide.

Lastly, this module will describe the imputation procedures associated with missing data within ELS:2002.

Slide 3 of 14

ELS is a sample that when appropriately weighted is nationally representative of students enrolled in 10th grade in the spring of 2002 or enrolled in the 12^{th} grade in the spring of 2004.

ELS is **not** a simple random sample (or SRS) of 10^{th} graders or 12^{th} graders, but rather a stratified, two-stage random sample for these two cohorts of students. For both cohorts, the school was selected at the first stage of sampling and students were randomly selected from those schools at the second stage. It's important to remember that ELS:2002 data cannot be analyzed as if they are SRS data. The standard assumption in basic inferential statistics (the assumption of simple random sampling) does not apply. Accordingly, weights must be used and standard errors must be appropriately calculated to ensure accurate analysis of ELS data.

Slide 4 of 14

As was discussed in the common module titled, 'Statistical Analysis of NCES Datasets Employing a Complex Design'**,** weights must be used in analyses of data from studies like ELS in order to make estimates produced from the sample representative of the target population. In ELS, the sampling weights are designed to make the sample representative of the population of students enrolled in 10th grade in the United States during the spring of 2002, and of students enrolled in the 12th grade in the spring of 2004.

As ELS is not a simple random sample, weights account for probability of selection and are an indicator of the relative importance of a particular observation to the computation of a sample statistic.

Additionally, not everyone sampled responded to ELS. Therefore, weights also account for differential response patterns within a given round (also referred to as wave or sometimes follow-up), as well as across time, to ensure that the data are representative of the population even though some members of the sample do not respond to the survey.

Ideally, there would be a nonresponse adjusted weight available for every component of every round of data collection. However, in longitudinal studies like ELS that have multiple components across multiple rounds of data collection, there are a number of possible weights that can be created for analysis of data within and across rounds. It is not economical, nor useful in a practical sense, to create nonresponse adjusted weights for every combination of components across every round of data collection. Therefore, each researcher will need to decide which of the weights provided is the best one to use, given a specific research question.

Slide 5 of 14

Before we talk about the types of weights available within ELS, let's revisit the purpose of ELS:2002 and consider the types of data available within ELS.

The first objective of ELS is to monitor the critical transitions made by a cohort of 2002 high school sophomores through college into their adult careers, with a special emphasis on college access and choice. The second objective is to obtain information about factors that influence these transitions, especially at the secondary level. Other ELS objectives include obtaining information about students and their educational experiences from multiple perspectives, as well as measuring gains in mathematics achievement from 10th to 12th grade, and maintaining comparability with other NCES secondary longitudinal studies.

As a result of these objectives, most ELS data can be classified as one of three types of variables: predictor variables, intermediate outcome variables, and outcome variables.

Almost all research questions that can be asked of ELS data will contain some combination of these three types of variables. Analysts may be interested in exploring the relationship between student characteristics and high school mathematics achievement. Others may be interested family background and college access. Initial postsecondary funding, including loan taking, might interest some researchers, with the outcome variable being timing of family formation.

In general, the ELS study was designed to collect information from students in early rounds of data collection that might be useful as predictor variables. Accordingly, most base year data describe student characteristics that were thought to be associated with later success in college and career. As the study progressed, later rounds of data collection focused on intermediate outcomes and additional contextual variables that might influence later outcomes. By the last data collections (F3 and postsecondary transcripts) the study sought mostly outcome data to document the college and career choices of study members.

To better frame the discussion regarding which weights to select when analyzing ELS data, we will consider two research questions regarding mathematics.

Slide 6 of 14

Let's now consider two types of research questions that can be asked of ELS data. In research question one, we are interested in the association between students' mathematics self-efficacy and achievement. In research question two, we are interested in if students' self-efficacy is sustained and then acted upon, as evidenced by declaring a mathematics-related postsecondary major within two years of high school completion.

Across both research questions, we will suppose a mathematics self-efficacy scale has been created using responses from the Base Year student questionnaire items including, "Because doing mathematics is fun, I wouldn't want to give it up", "Mathematics is important to me personally", and "I'm certain I can master the skills being taught in my math class".

To address the first research question, analysts would use the base year mathematics self-efficacy related variables and variables from the base year administration of the ELS mathematics assessment. As both of these sources of data were collected during the same data collection round – the base year – selecting and using a base year weight to conduct the analysis would be most appropriate.

However, in the case of the second research question data are needed from multiple rounds of data collection to properly address the question. The base year self-efficacy variables could be considered again to address part of this research question. However, we would also need to examine the degree to which students' reported comparable responses to the item "I'm certain I can master the skills being taught in my math class" across both the base year and the first follow-up student questionnaires. Additionally, we would need to examine the responses that students provided when asked to indicate their major field of study within the second follow-up student questionnaire. Selecting a weight for use in this analysis becomes a bit more challenging.

Slide 7 of 14

taken in part from the base year to third follow-up data file documentation, shows a few of those weights and their relationship to survey respondents. The most important thing to keep in mind is that there are essentially four sets of weights, one from each round of data collection. Here you see two base year weights, five first follow-up weights, four second follow-up weights, and three third follow-up weights.

To address research question one, “Is there an association between students’ mathematics self-efficacy and mathematics achievement?” analysts would need to select BYSTUWT, spelled B-Y-S-T-U-W-T, to conduct their analysis. BYSTUWT is a cross-sectional weight as it pertains to only one round of data collection. Other cross-sectional weights listed here are F1QWT (spelled F-1-Q-W-T), F2QWT, and F3QWT.

These cross-sectional weights are only useful when analyzing data from the specified round. For example, a base year student weight is no longer useful when analyzing data from the first follow-up or any other data collection round. Be sure to select the cross-sectional weight that corresponds to the round of data from which your variables of interest originate.

To address research question two, “Do students who report mathematics self-efficacy in 10th grade sustain that self-efficacy in the 12th grade and then declare a mathematics-related postsecondary major field of study within two years of high school completion?” analysts would need to select a panel weight. Panel weights are used when your analysis spans more than one round of data collection. Panel weights can be created for any combination of responses by round. In this table the panel weights are F1PNLWT (spelled F-1-P-N-L-W-T), F1TRSCWT, F2QTSCWT, F2F1WT, F2BYWT, F3QTSCWT, and F3BYPNLWT (spelled F-3-B-Y-P-N-L-W-T).

In the case of research question two, the panel weight F2BYWT would be the most appropriate weight to select to address the question as the predictor or independent variable (mathematics self-efficacy – measured in the base year) and the outcome or dependent variable (mathematics major – measured in the second follow-up) are found in the two rounds of data collection associated with this weight.

Remember that you should choose the weight that most closely matches the rounds of data associated with the independent and dependent variables included in your analysis. In this way, your analytical results will be least affected by survey nonresponse.

Slide 8 of 14

Both cross-sectional and panel weights are created by the same process of non-response adjustment which includes distributing the design weights of survey non-respondents to survey respondents to maintain the national representativeness and total population weight of the sample.

Cross-sectional and panel weights are created by distributing the weights of non-respondents within one round of the survey, such as the base year, to others in the round who are ‘similar’ to those who responded in terms of key frame variables which included enrollment status, school sector, region, and percent of students on free or reduced-price lunch, percent of certified teachers, percent of students with IEPs, and race/ethnicity. These frame variables varied across ELS data collection rounds and are described in detail within the technical reports that accompany each round of ELS data.

The primary reason this “non-response adjustment” is necessary is that members of the sample who did not respond may disproportionately come from subgroups of the sample who were sufficiently different from those who responded. If this were the case, the data collected would no longer be nationally representative. Accordingly, this potential for distortion (or bias) was removed by adjusting the design weights to account for sample member nonresponse.

Slide 9 of 14

Data from each of the follow-up ELS data collections include both the original sophomore cohort and the freshened senior cohort. Accordingly, cohort flags are needed to be able to make statements about a nationally representative sample of either sophomores or seniors. Cohort flags are found within the ID and Universe Variables student file within eDAT and within the ECB on the restricted-use CD/DVD. The cross-sectional analysis weights F1QWT (spelled F-1-Q-W-T), F2QWT and F3QWT were created by the process of adjustment described on the previous screen – that is, the design weights of the first follow-up nonrespondents were distributed to first follow-up respondents to maintain the national representativeness and total population weight of the sample.

The design weight in this process was essentially the analysis weight from the previous round of data collection **with** the added senior sample members, whose design weights correspond to their probabilities of selection.

The adjustment process results in a first follow-up (or F1) cross-sectional analysis weight, called F1QWT, which includes members of both the 2002 sophomore and 2004 senior groups. Consequently, the straight application of this cross-sectional weight to the ELS sample would produce incorrect results. Analysts must use a cohort flag with all follow-up cross-sectional analysis weights.

Slide 10 of 14

In the common module titled, ‘Statistical Analysis of NCES Datasets Employing a Complex Sample Design’, two standard error calculation procedures were discussed: Replication Techniques and Taylor Series linearization. This portion of the ELS training modules describes important information regarding the inclusion of study design variables in your statistical package to calculate standard errors.

Replication is a method that calculates appropriate standard errors based on differences between estimates from the full sample and a series of created subsamples, or replicates. If this is the method you will use to calculate standard errors, you need to select replicate weights that are associated with your main sampling weight. Every main sampling weight in ELS has a set of replicate weights that are associated with it. For example, the replicate weights associated with the base year school cross-sectional weight BYSCHWT (spelled B-Y-S-C-H-W-T) are BYSCH1 through BYSCH200 (spelled B-Y-S-C-H-1 and B-Y-S-C-H-200, respectively).

ELS:2002 replication weights use a Balanced Repeated Replication method so that is the method that should be specified within the statistical software used for analysis.

Slide 11 of 14

The second method for calculating standard errors is the Taylor series linearization method. This method uses PSU and strata identifiers to compute the appropriate standard errors. This method of standard error calculation is only available to users with access to restricted-use data, as the stratum and PSU variables are not available on the public-use file to minimize disclosure risk. If this is the method you will use to compute standard errors, you will select the PSU and strata identifiers associated with your sampling weight named ‘PSU’ and ‘STRAT_ID’(spelled S-T-R-A-T-underscore-I-D), and specify Taylor Series linearization as the method within the statistical software used for analysis.

Slide 12 of 14

Missing data in ELS are handled in two ways.

First, most of the Base Year non-respondents returned to the sample in the First Follow-up and, if they responded, most of the student background data that would have been collected from them had they responded in the base year were collected from them in the first follow-up. However, nonrespondents to both the base year and first follow-up were removed from the sample before the second follow-up.

Item-level non-response in ELS regarding student background variables was handled through imputation, meaning all missing data in the base year and first follow-up student background variables were imputed. Therefore, all student background composites have values for all sample members.

Imputation flags are used in ELS whenever variables were imputed. For example, the student’s sex, collected in the first follow-up, was imputed if not known. Variable F1SEXIM (spelled F-1-S-E-X-I-M) indicates whether a given student’s information for F1SEX was imputed. In this way, researchers may decide whether to use imputed or raw data within a specific analysis.

Slide 13 of 14

This table provides an example of how the imputation process works in ELS for the gender variables BYS14, BYSEX, and F1SEX.

BYS14 is labeled ‘questionnaire’ because it includes only questionnaire responses provided by sample members. As you can see in the table, 15,183 sample members provided a positive response value of either male {1} or female {2}. 1,014 respondents are coded using a negative response value (or missing data/reserve code) which provides insight into why the gender variable is missing.

The largest type of missing data comes from base year "Nonrespondents," who are coded using {-4}. These 648 sample members who could not be located or never responded to the base year request to complete a questionnaire interview, were subsequently returned to the sample during the first follow-up.

There are 305 students coded as {-8} or legitimate skip/NA who include freshened students and base year incapable students. There are also 59 students coded as "Missing" or {-9} who were students who for one reason or another didn't answer the question that asked them about their gender. Students coded as {-2}, or "Refused", which included only two sample members for this question, represents sample members who refused to answer the question.

For the base year data release, imputations were done for three of the four groups of missing data/reserve codes, 'refusals,' 'legitimate skip/NA,' and 'missing,' to produce a composite called BYSEX. The imputations were done, first, "logically" and then, if logical imputation could not be done, "statistically."

The most important source of information for these "logical imputations" was the school roster. In some cases gender was inferred from the student's name as shown on the school roster or taken directly from the available school roster information.

Then, in the First Follow-up data collection, attempts were made to fill-in all the remaining negative response values for the group of base year nonrespondents by returning them to the sample. Returning students were administered the same basic student background questions that were on the Base Year Student Questionnaire, facilitating imputation.

As a result of these steps, gender is provided in the F1SEX variable for all of the base year and first follow-up sample members, including all those who responded in base year and those who didn't respond until returned to the sample in the first follow-up.

Consequently, because the variable F1SEX has no missing data, it should be used in all analyses. As all base year and first follow-up composite variables on the ELS data files have been imputed, analysts should use these composite variables in place of the related raw data in all analyses.

Slide 14 of 14

This module presented information about the ELS survey design to facilitate a more detailed discussion of sampling weights, standard error calculation, and missing and imputed data. This module described the types of weights within ELS:2002 and provided guidance regarding how to select the appropriate weight for specific analyses.

Additionally, important resources that have been provided throughout the module are summarized here along with the module's objectives for your reference.

You may now proceed to the next module in the series or exit the module.